



GC830-PCT.txt  
SEQUENCE LISTING

<110> Genencor International, Inc.  
Dunn-Coleman, N.  
Shetty, J.  
Duan, G.  
Sung, A.  
Qian, Y.

<120> A Method for the Preparation of a High Purity Rice Protein Concentrate

<130> GC830-PCT

<140> PCT/US05/05320  
<141> 2005-02-17

<150> US 60/547,153  
<151> 2004-02-23

<160> 11

<170> PatentIn version 3.2

<210> 1  
<211> 2103  
<212> DNA  
<213> Humicola grisea var. thermoidea

<400> 1

atgcataacct	tctccaagct	cctcgccctg	ggctctgccg	tccagtctgc	cctcggcg	60
cctcaccgct	cttcgcgtct	ccaggaacgc	gctgccgttg	ataccttcat	caacaccgag	120
aagcccacatcg	catggaaacaa	gctgctcgcc	aacatcgcc	ctaacggcaa	agccgctccc	180
ggtgccgccc	ccggcggtgt	gattgccagc	cttccagga	cggaccctcc	ttgtacgtgg	240
tggcatggaa	tggacccaag	agactggttt	tagatgaaag	agagttctg	ctaaccgcca	300
cacccagact	tcttcacctg	gaccgcgat	gccgcctgg	tcctcaccgg	catcatcgag	360
tcccttggcc	acaactacaa	caccaccctg	cagaccgtca	tccagaacta	cgtcgcgtcg	420
caggccaaagc	tgcagcaggt	ctcgaacccc	tcggaaacct	tcgcccacgg	ctcggtctc	480
ggtgaggcca	agttcaatgt	cgacccact	gccttcactg	gcgaatgggg	tgccttcag	540
agggacggcc	cggccctcgcg	cgccatcgct	ctcatccagt	acgccaagtg	gctgatcgcc	600
aacggctaca	agagcacggc	caagagcgtc	gtctggccc	tcgtcaagaa	cgatctcgcc	660
tacacggccc	agtactggaa	cgagaccggc	ttcgatctt	gggaggaggt	ccccggcagc	720
tcgttcttta	ccatcgccag	ctctcacagg	ggtgagtcat	ttattgttca	gtgtttctc	780
attgaataat	taccggaaatg	ccactgacgc	caaacagctc	tgactgaggg	tgcttacctc	840
gccgctcagc	tcgacacccga	gtgcccggcc	tgcacgaccg	tcgcccctca	ggttctgtgc	900
ttccagcagg	ccttctggaa	ctccaagggc	aactatgtcg	tctccaacag	taagatccct	960
acaccaacaa	aaaaaatcga	aaaggaacgt	tagctgaccc	ttcttagtcaa	cggcggcgag	1020
tatcgctccg	gcaaggacgc	caactcgatc	ctggcgtcca	tccacaactt	cgaccctgag	1080
gccggctcg	acaacactgac	ttccagccc	tgcagcggac	gcccctggc	caaccacaag	1140
gcctatgtcg	actcggtccg	caacctctac	gccatcaaca	agggcatcg	ccagggcaag	1200
gccgttggcc	tcggccgcta	ctcgaggat	gtctactaca	acggcaaccc	gtggtacctg	1260
gccaactttg	ccggccggca	gcagctctac	gacgccatct	acgtgtggaa	caagcaggc	1320
tccatcaccc	tgcacccgt	ctccctgccc	ttcttcgctg	accttgtctc	gtcggtcagc	1380
accggcacct	actccaagag	cagctcgacc	ttcaccaaca	tcgtcaacgc	cgtcaaggcc	1440
tacggccacg	gcttcatcg	ggtggcgccc	aagtacaccc	cgtccaacgg	cgcgctcgcc	1500
gagcagtagc	accgcaacac	ggcaagccc	gactcgccg	ccgacccgtac	gtggtctac	1560
tcggccctcc	tctcgccat	cgaccggccg	gccccgtctg	tcccccggag	ctggcgggccc	1620
agcgtggcca	agagccagct	gcccgtccacc	tgctcgcc	tcgagggtcg	cggcacccat	1680
gtcgccgcca	cgagcaccc	gttcccgatcc	aaggcagaccc	cgaaccctc	cgcggcgccc	1740

## GC830-PCT.txt

tccccgtccc	cctacccgac	cgcctgcg	gacgctagcg	agggtgtacgt	caccttaac	1800
gagcgcgtgt	cgaccgcgtg	ggcgagacc	atcaagggtgg	tgggcaacgt	gccggcgctg	1860
gggaactggg	acacgtccaa	ggcggtgacc	ctgtcgccca	gcgggtacaa	gtcgaatgat	1920
cccctcttgg	gcatcacgg	gcccataag	gacgacgggct	cggccgtgca	gtacaagtt	1980
atcaagggtcg	gcaccaacgg	gaagattact	tgggagtcgg	accccaacag	gagcattacc	2040
ctgcagacgg	cgtcgcttc	ggcaagtgc	gccgcgcaga	cggtgaatga	ttcgtggcgt	2100
taa						2103

<210> 2  
<211> 634  
<212> PRT  
<213> Humicola grisea var. thermoidea

<400> 2						
Met His Thr Phe Ser Lys Leu Leu Val	Leu Gly Ser Ala Val Gln Ser					
1	5	10	15			
Ala Leu Gly Arg Pro His Gly Ser Ser	Arg Leu Gln Glu Arg Ala Ala					
20	25	30				
Val Asp Thr Phe Ile Asn Thr Glu Lys	Pro Ile Ala Trp Asn Lys Leu					
35	40	45				
Leu Ala Asn Ile Gly Pro Asn Gly Lys	Ala Ala Pro Gly Ala Ala Ala					
50	55	60				
Gly Val Val Ile Ala Ser Pro Ser Arg	Thr Asp Pro Pro Tyr Phe Phe					
65	70	75	80			
Thr Trp Thr Arg Asp Ala Ala Leu Val	Leu Thr Gly Ile Ile Glu Ser					
85	90	95				
Leu Gly His Asn Tyr Asn Thr Thr Leu	Gln Thr Val Ile Gln Asn Tyr					
100	105	110				
Val Ala Ser Gln Ala Lys Leu Gln Gln	Val Ser Asn Pro Ser Gly Thr					
115	120	125				
Phe Ala Asp Gly Ser Gly Leu Gly Glu	Ala Lys Phe Asn Val Asp Leu					
130	135	140				
Thr Ala Phe Thr Gly Glu Trp Gly Arg	Pro Gln Arg Asp Gly Pro Pro					
145	150	155	160			
Leu Arg Ala Ile Ala Leu Ile Gln Tyr	Ala Lys Trp Leu Ile Ala Asn					
165	170	175				
Gly Tyr Lys Ser Thr Ala Lys Ser Val	Val Trp Pro Val Val Lys Asn					
180	185	190				
Asp Leu Ala Tyr Thr Ala Gln Tyr Trp	Asn Glu Thr Gly Phe Asp Leu					
195	200	205				
Trp Glu Glu Val Pro Gly Ser Ser Phe	Phe Thr Ile Ala Ser Ser His					
210	215	220				
Arg Ala Leu Thr Glu Gly Ala Tyr Leu	Ala Ala Gln Leu Asp Thr Glu					
225	230	235	240			
Cys Arg Ala Cys Thr Thr Val Ala Pro	Gln Val Leu Cys Phe Gln Gln					
245	250	255				
Ala Phe Trp Asn Ser Lys Gly Asn Tyr	Val Val Ser Asn Ile Asn Gly					
260	265	270				
Gly Glu Tyr Arg Ser Gly Lys Asp Ala	Asn Ser Ile Leu Ala Ser Ile					
275	280	285				
His Asn Phe Asp Pro Glu Ala Gly Cys	Asp Asn Leu Thr Phe Gln Pro					
290	295	300				
Cys Ser Glu Arg Ala Leu Ala Asn His	Lys Ala Tyr Val Asp Ser Phe					
305	310	315	320			
Arg Asn Leu Tyr Ala Ile Asn Lys Gly	Ile Ala Gln Gly Lys Ala Val					
325	330	335				
Ala Val Gly Arg Tyr Ser Glu Asp Val	Tyr Tyr Asn Gly Asn Pro Trp					
340	345	350				
Tyr Leu Ala Asn Phe Ala Ala Glu Gln	Leu Tyr Asp Ala Ile Tyr					

## GC830-PCT.txt

355	360	365
Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser Val Ser Leu Pro		
370	375	380
Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr Ser Lys		
385	390	395 400
Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val Lys Ala Tyr Ala		
405	410	415
Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro Ser Asn Gly Ala		
420	425	430
Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro Asp Ser Ala Ala		
435	440	445
Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala Ile Asp Arg Arg		
450	455	460
Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val Ala Lys Ser Gln		
465	470	475 480
Leu Pro Ser Thr Cys Ser Arg Ile Glu Val Ala Gly Thr Tyr Val Ala		
485	490	495
Ala Thr Ser Thr Ser Phe Pro Ser Lys Gln Thr Pro Asn Pro Ser Ala		
500	505	510
Ala Pro Ser Pro Ser Pro Tyr Pro Thr Ala Cys Ala Asp Ala Ser Glu		
515	520	525
Val Tyr Val Thr Phe Asn Glu Arg Val Ser Thr Ala Trp Gly Glu Thr		
530	535	540
Ile Lys Val Val Gly Asn Val Pro Ala Leu Gly Asn Trp Asp Thr Ser		
545	550	555 560
Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr Lys Ser Asn Asp Pro Leu		
565	570	575
Trp Ser Ile Thr Val Pro Ile Lys Ala Thr Gly Ser Ala Val Gln Tyr		
580	585	590
Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys Ile Thr Trp Glu Ser Asp		
595	600	605
Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala Ser Ser Ala Gly Lys Cys		
610	615	620
Ala Ala Gln Thr Val Asn Asp Ser Trp Arg		
625	630	

<210> 3  
<211> 604  
<212> PRT  
<213> Humicola grisea var. thermoidea

<400> 3		
Ala Ala Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn		
1	5	10 15
Lys Leu Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala		
20	25	30
Ala Ala Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr		
35	40	45
Phe Phe Thr Trp Thr Arg Asp Ala Ala Leu Val Leu Thr Gly Ile Ile		
50	55	60
Glu Ser Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Thr Val Ile Gln		
65	70	75 80
Asn Tyr Val Ala Ser Gln Ala Lys Leu Gln Gln Val Ser Asn Pro Ser		
85	90	95
Gly Thr Phe Ala Asp Gly Ser Gly Leu Gly Glu Ala Lys Phe Asn Val		
100	105	110
Asp Leu Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg Asp Gly		
115	120	125
Pro Pro Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp Leu Ile		

## GC830-PCT.txt

130	135	140
Ala Asn Gly Tyr Lys Ser Thr Ala Lys Ser Val Val Trp Pro Val Val		
145	150	155 160
Lys Asn Asp Leu Ala Tyr Thr Ala Gln Tyr Trp Asn Glu Thr Gly Phe		
165	170	175
Asp Leu Trp Glu Glu Val Pro Gly Ser Ser Phe Phe Thr Ile Ala Ser		
180	185	190
Ser His Arg Ala Leu Thr Glu Gly Ala Tyr Leu Ala Ala Gln Leu Asp		
195	200	205
Thr Glu Cys Arg Ala Cys Thr Thr Val Ala Pro Gln Val Leu Cys Phe		
210	215	220
Gln Gln Ala Phe Trp Asn Ser Lys Gly Asn Tyr Val Val Ser Asn Ile		
225	230	235 240
Asn Gly Gly Glu Tyr Arg Ser Gly Lys Asp Ala Asn Ser Ile Leu Ala		
245	250	255
Ser Ile His Asn Phe Asp Pro Glu Ala Gly Cys Asp Asn Leu Thr Phe		
260	265	270
Gln Pro Cys Ser Glu Arg Ala Leu Ala Asn His Lys Ala Tyr Val Asp		
275	280	285
Ser Phe Arg Asn Leu Tyr Ala Ile Asn Lys Gly Ile Ala Gln Gly Lys		
290	295	300
Ala Val Ala Val Gly Arg Tyr Ser Glu Asp Val Tyr Tyr Asn Gly Asn		
305	310	315 320
Pro Trp Tyr Leu Ala Asn Phe Ala Ala Glu Gln Leu Tyr Asp Ala		
325	330	335
Ile Tyr Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser Val Ser		
340	345	350
Leu Pro Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr		
355	360	365
Ser Lys Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val Lys Ala		
370	375	380
Tyr Ala Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro Ser Asn		
385	390	395 400
Gly Ala Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro Asp Ser		
405	410	415
Ala Ala Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala Ile Asp		
420	425	430
Arg Arg Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val Ala Lys		
435	440	445
Ser Gln Leu Pro Ser Thr Cys Ser Arg Ile Glu Val Ala Gly Thr Tyr		
450	455	460
Val Ala Ala Thr Ser Thr Ser Phe Pro Ser Lys Gln Thr Pro Asn Pro		
465	470	475 480
Ser Ala Ala Pro Ser Pro Ser Pro Tyr Pro Thr Ala Cys Ala Asp Ala		
485	490	495
Ser Glu Val Tyr Val Thr Phe Asn Glu Arg Val Ser Thr Ala Trp Gly		
500	505	510
Glu Thr Ile Lys Val Val Gly Asn Val Pro Ala Leu Gly Asn Trp Asp		
515	520	525
Thr Ser Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr Lys Ser Asn Asp		
530	535	540
Pro Leu Trp Ser Ile Thr Val Pro Ile Lys Ala Thr Gly Ser Ala Val		
545	550	555 560
Gln Tyr Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys Ile Thr Trp Glu		
565	570	575
Ser Asp Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala Ser Ser Ala Gly		
580	585	590
Lys Cys Ala Ala Gln Thr Val Asn Asp Ser Trp Arg		
595	600	

## GC830-PCT.txt

<210> 4  
 <211> 10739  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> pTrex3g\_N13 plasmid  
  
 <400> 4

aagcttacta	gtacttctcg	agctctgtac	atgtccggtc	gcgcacgtacg	cgtatcgatg	60
gcccgcgt	caggcggccg	cctgcagcca	cttgcagtcc	cgtgaaattc	tcacggtgaa	120
tgtaggcctt	ttttagggta	ggaattgtca	ctcaagcacc	cccaacctcc	attacgcctc	180
ccccatagag	ttcccaatca	gtgagtcatg	gcactgttct	caaatacgatt	ggggagaagt	240
tgacttccgc	ccagagctga	aggtcgacaca	accgcatgat	atagggtcgg	caacggcaaa	300
aaagcacgtg	gctcaccgaa	aagcaagatg	tttgcgatct	aacatccagg	aacctggata	360
catccatcat	cacgcacgac	cacttgcattc	tgctggtaaa	ctcgatttcg	ccctaaaccg	420
aagtgcgtgg	taaatctaca	cgtggccccc	tttgcgtata	ctgcgtgtgt	cttctctagg	480
tgccatctt	ttcccttcct	ctagttgtga	attgtttgtg	ttggagtcgg	agctgtact	540
acctctgaat	ctctggagaa	tggtgacta	acgactaccc	tgcacctgca	tcatgtatata	600
aatagtgtatc	ctgagaaggg	gggttggag	caatgtggg	ctttgatgtt	catcaaacaa	660
agaacgaaga	cgcctctttt	gcaaaagtttt	gtttcggcta	cggtaagaa	ctggataactt	720
gttgtgtctt	ctgtgtattt	tttgtgcaac	aagaggccag	agacaatcta	ttcaaaccacc	780
aagcttgctc	ttttgagcta	caagaacctg	ttgggtatata	atctagagtt	gtgaagtcgg	840
taatcccgct	gtatagtata	acgactcgca	tctaaatact	ccgaagctgc	tgcaacccg	900
gagaatcgag	atgtgcgtga	aagcttctag	cgagcggcta	aattagcatg	aaaggctatg	960
agaaattctg	gagacggctt	gttgaatcat	gcccgttccat	tcttcgacaa	gcaaaagcgtt	1020
ccgtcgca	agcaggcact	cattcccgaa	aaaactcgga	gattcctaag	tagcgatgga	1080
accggaataa	tataataggc	aatacattga	gttgcctcga	cggttgcaat	gcaggggtac	1140
ttagcttgg	cataactgtt	ccgtacccca	cctttctca	acctttggcg	tttccctgtat	1200
tcagcgtacc	cgtacaagtc	gtaatacta	ttaacccaga	ctgaccggac	gtgtttgcc	1260
tttcatttgg	agaaaataatg	tcatttgcgt	gtgttaattt	ctgtgttgac	cgactggggc	1320
tgttcgaagc	ccgaatgtag	gatttttattc	cgaactctgc	tcgtagaggc	atgttgtgaa	1380
tctgtgtcgg	gcaggacacg	cctcgaaggt	tcacggcaag	ggaaaccacc	gatagcaatg	1440
tctagtagca	acctgtaaag	ccgcaatgca	gcatcactgg	aaaataaaaa	ccatggcta	1500
aaagtacata	atgttatgtcc	taaagaagtc	atataccacg	ggctataata	tgtacaatca	1560
agtggctaaa	cgtaccgtaa	tttgccaaacg	gcttgggggg	ttgcagaagc	aacggcaaa	1620
ccccacttcc	ccacgtttgt	tttttcactc	agtccaaatct	cagctggta	tcccccatt	1680
gggtcgcttg	tttgcgtccgg	tgaagtgaaa	gaagacagag	gtaagaatgt	ctgactcgga	1740
gcgttttgc	tacaaccaag	ggcagtgtat	gaagacagtg	aatatgttgc	attcaaggag	1800
tatatttgc	gggatgttttgc	agtgtatctgt	gtaaggaggt	ttgtctggcg	atacgacgaa	1860
tactgtatag	tcacttctga	tgaagtggtc	catattgaaa	tgtaaatgtc	gcactgaaca	1920
ggcaaaagat	ttagttgaaa	ctgcctaaga	tctcgggccc	tcgggccttc	ggccctttggg	1980
tgtacatgtt	tgtgtccgg	gcaaatgca	agtgtggtag	gatcgaaacac	actgtctgcct	2040
tttaccaagca	gtgtgggtta	tgtgtatggc	aatgttca	gggcacttc	atggtttgcg	2100
atagaaagag	aagcttagcc	aagaacaata	gccgataaa	atagcctcat	taaacggaa	2160
gagcttagtag	gcaaagtca	cgaatgtgt	tatataaaagg	ttcgagggtcc	gtgcctccct	2220
catgtctcc	ccatctactc	atcaactca	atcctccagg	agacttgtac	accatcttt	2280
gaggcacacaga	aacccaata	tcaaccatca	caagtttgc	caaaaaagca	ggctccgcgg	2340
ccgccccctt	caacatgc	accttctcca	agctcctcgt	cctggctct	gccgtccagt	2400
ctgccttcgg	ggccgcctc	ggctttcgc	gtctccagga	acgcgcgtcc	gttgataacct	2460
tcatcaacac	cgagaagccc	atcgcatgga	acaagctgt	cgccaaacatc	ggcccta	2520
gcaaagccgc	tcccggtgc	ggccgcggcg	ttgtgattgc	cagccctcc	aggacggacc	2580
ctccttgtac	gtgggtggcat	ggaatggacc	caagagactg	gttttagatg	aaagagagtt	2640
tctgtctacc	gccacaccca	gacttctca	cctggacc	cgatgcc	ctggccctca	2700
ccggcatcat	cgagtccctt	ggccacaact	acaacaccac	cctgcagacc	gtcatccaga	2760
actacgtcgc	gtcgccaggcc	aagctgc	aggctcgaa	cccctcg	accttcgcgg	2820
acggctcg	tctcggtgag	gccaagttca	atgtcgact	cactgcctc	actggcgaat	2880
ggggtcgccc	tcagagggac	ggccccc	tgcgc	cgctctc	cagtacgcca	2940

## GC830-PCT.txt

agtggctgat	cgcacacggc	tacaagagca	cggccaagag	cgtcgctgg	cccgtcgta	3000
agaacgatct	cgcctacacg	gcccagtact	ggaacgagac	cggcttcgat	ctctggagg	3060
aggccccgg	cagctcgttc	tttaccatcg	ccagctctca	caggggttag	tcattttatg	3120
ttcagtgttt	tctcattgaa	taattaccgg	aatgccactg	acgccaaaca	gctctgactg	3180
agggtgctta	cctcgccgct	cagctcgaca	ccgagtgccg	cgcctgcacg	accgtcgccc	3240
ctcaggttct	gtgcttccag	caggccttct	ggaactccaa	ggcaactat	gtcgctcca	3300
acagaat	ccctacacca	acaaaaaaaaa	tcgaaaagg	acgttagctg	acccttctag	3360
tcaacgcgg	cgagtatcgc	tccggcaagg	acgccaactc	gatcctggcg	tccatccaca	3420
acttcgaccc	tgaggccggc	tgcgacaacc	tgacccctca	gcccgtcage	gagcgcgccc	3480
tggccaacca	caaggcctat	gtcgactctg	tccgcaacat	ctacgcccac	aacaagggca	3540
tcgcccagg	caaggccgtt	gccgtcgcc	gctactcgga	ggatgtctac	tacaacggca	3600
accctgtgtt	cctggccaac	tttgcggcc	ccgagcagct	ctacgacgac	atctacgtgt	3660
ggaacaaga	gggctccatc	accgtgaccc	cggtctccct	gccccttctc	cgcgacccct	3720
tctcgtcggt	cagcacccggc	acctactcca	agagcagctc	gaccccttacc	aacatcgta	3780
acgcccataa	ggcctacgac	gacggcttca	tcgaggtggc	ggccaagttac	acccctgtcca	3840
acggcgcgt	cgccgagcag	tacgaccgca	acacggggca	gcccactcg	gcccgcgacc	3900
tgacgtggc	gtactcgcc	ttcctctcg	ccatcgaccg	ccgcgcgggt	ctcgcccccc	3960
cgagctggcg	ggccagcgtt	gccaagagcc	agctgcccgt	cacctgctcg	cgcacatcgagg	4020
tcgcccgcac	ctacgtcgcc	gccacggca	cctcgcttcc	gtccaagcag	accccgaaacc	4080
cctccgcggc	gccctcccccc	tccccctacc	cgaccgcctg	cgcgacgct	agcgaggtgt	4140
acgtcacctt	caacgagcgc	gtgtcgaccg	cgtggggcga	gaccatcaag	gtgtgggca	4200
acgtgcggc	gctggggaaac	tgggacacgt	ccaaggcgtt	gaccctgtcg	gccagcgggt	4260
acaagtcgaa	tgtatcccctc	tggagcatca	cggtgcccatt	caagcgcacg	ggctcgcccg	4320
tgcagtacaa	gtatatacg	gtcggcacca	acgggaagat	tacttggag	tcggaccctt	4380
acaggagcat	taccctgcag	acggcgctgt	ctgcgggcaa	gtgcgcgcg	cagacgggt	4440
atgattcgtg	gctttaaaag	ggtgggcgcg	ccgaccacagc	tttcttgcac	aaagtgggt	4500
tcgcgcac	tccgtcgaa	agcctgacgc	accggtagat	tcttggtag	cccgatcat	4560
gacggcggcg	ggagctacat	ggccccgggt	gatttatttt	ttttgtatct	acttctgacc	4620
cttttcaaata	atacgtca	ctcatcttc	actggagatg	ccgcctgcctt	ggtattgcga	4680
tgttgtcagc	ttggcaattt	gtggcttgc	aaaacacaaa	acgattccctt	atgtgcatt	4740
cattttaa	taacgaaata	gaagaaagag	gaaattaaaa	aaaaaaaaaa	aacaaacatc	4800
ccgttcataa	cccgtagaaat	cgccgtctt	cgtgtatccc	agtaccagtt	tattttgaat	4860
agctcgcccg	ctggagagca	tcctgaatgc	aagtaacaac	cgtagaggct	gacacggcag	4920
gtgttgcgt	ggagcgtcgt	gttctacaag	gccagacgtc	ttcgcgggtt	atatatatgt	4980
atgtttgact	gcaggcgtct	cagcgtacgc	agtcaagttc	gcccctcgctg	cttgcataat	5040
aatcgcaat	gggaagccac	accgtgactc	ccatcttca	gtaaagctt	gttgggtttt	5100
atcagcaata	cacgttaattt	aaactcgta	gcatggggct	gatagctta	ttaccgttta	5160
ccagtgccat	ggttctgcag	cttccttgg	cccgtaaaaat	tcggcgaagc	cagccaatca	5220
ccagctaggc	accagctaaa	ccctataatt	agtctcttat	caacaccatc	cgctcccccg	5280
ggatcaatga	ggagaatgag	ggggatgcgg	ggctaaagaa	gcctacataa	ccctcatgccc	5340
aactcccagt	ttacactcg	cgagccaaca	tcctgactat	aagctaaca	agaatgcctc	5400
aatcctggg	agaactggcc	gctgataagc	gcccggccct	cgcaaaaaacc	atccctgtat	5460
aatggaaagt	ccagacgctg	cctgcggaa	acagcgttat	tgatttccca	aagaaatcggt	5520
ggatccttcc	agaggccgaa	ctgaagatca	cagaggcctc	cgctgcagat	cttgcgttca	5580
agctggcgcc	cggagagttt	acctcggtt	aaagtacgt	agcattctgt	aaacgggcag	5640
caatcgccca	cgagttagta	gggtcccttc	tacctctcg	ggagatgtaa	caacgcacc	5700
ttatggact	atcaagctga	cgctggctt	tgtgcagaca	aactgcgtcc	acgagtttctt	5760
ccctgacgccc	gctctcgcc	aggcaaggga	actcgatgaa	tactacgca	agcacaagag	5820
acccgttgg	ccactccat	gcctcccat	ctctctcaaa	gaccagctt	gagtcaaggt	5880
acaccgttgc	ccctaagtcg	ttagatgtcc	ctttttgtca	gctaacat	gccaccagg	5940
ctacgaaaca	tcaatggct	acatctcat	gctaaacaag	tacgacgaag	gggactcggt	6000
tctgacaacc	atgctccgca	aagccggc	cgtttctac	gtcaagac	ctgtcccgca	6060
gaccctgt	gtctgcgaga	cagtcaacaa	catcatcg	cgcaccgtca	acccacgca	6120
caagaactgg	tcgtcgccg	gcagttctt	ttgtgagggt	gcgatcg	ggattcgcrv	6180
tggtggcg	atcggtgt	gaacggat	cggtggctcg	attcgatgtc	cggccgcgtt	6240
caacttcc	tacggctaa	ggccgagtca	ttggcggctg	ccgtatgca	agatggcgaa	6300
cacgtggag	ggtcaggaga	cggtgacacg	cggtgtcggg	ccgattacgc	actctgttga	6360
gggtgagtcc	ttcgcctt	ccttcttcc	ctgctctata	ccaggcctcc	actgtcctcc	6420
tttcttgc	tttatactat	atacgagacc	ggcagtcact	gatgaagtat	gttagaccc	6480

## GC830-PCT.txt

cgcctttca ccaaatccgt cctcggtcag gagccatgaa aatacgactc caaggcatc	6540
cccattcccc ggcgccagtc cgagtcggac attattgcct ccaagatcaa gaacggcggg	6600
ctcaaatatcg gctactcaa ctgcgacggc aatgtcctt cacaccctcc tatectgcgc	6660
ggcgtggaaa ccaccgtcgc cgcactcgcc aaagccggc acaccgtac cccgtggacg	6720
ccataacaagg acgatttcgg ccacgatctc atctccata tctacgcggc tgacggcagc	6780
crvgccgacg taatgcgcga tatcagtgcg tccggcgagc cggcattcc aatatcaaa	6840
gacctactga acccgAACat caaagctgtt aacatgaacg agctctggg cacgcacatc	6900
cagaagtggg attaccagat ggagtaccc gagaatggc gggaggctga agaaaaggcc	6960
gggaagggAAC tggacggccat catcgccgcg attacgccta ccgcgtcggt acggcatgac	7020
cagttccgg actatgggtt tgcctctgtg atcaacctgc tggatttac gagcgtgggt	7080
gttccggta cctttgcggta taagaacatc gataagaaga atgagagttt caaggcgggt	7140
agttagctt atgcctcgat gcaggaagag tatgatccgg aggcttacca tggggcaccc	7200
gttgcagtgc aggttatcg acggagactc agtgaagaga ggacgttggc gattgcagag	7260
gaagtgggg agttgttggg aaatgtgggt actccatagc taataagtgt cagatagcaa	7320
tttgacacaag aaatcaatac cagcaactgt aaataagcgc tgaagtgacc atgcacatgt	7380
acgaaagagc agaaaaaaac ctgcccgtaga accgaagaga tatgacacgc ttccatctct	7440
caaaggaaga atcccttcag gttgcgttt ccagtctaga cacgtataac ggcacaatgt	7500
tctctcacca aatgggttat atctcaaatg tgcgttcaagg atggaaagcc cagaatatcg	7560
atcgccgcga gatccatata tagggcccg gttataatata cctcaggctc acgtccatcg	7620
gccattcgaa ttctgtatca tggctatagc tggttcctgt gtgaaattgt tatccgtca	7680
caattccaca caacatacga gccggaaagca taaagtgtaa agcctgggt gcctaattgag	7740
tgagctaact cacattaatt gcgttgcgt cactgcccgc ttccagtcg gaaaacctgt	7800
cgtgccagct gcattaatga atcggccaac ggcggggag aggccgttt cgtattggc	7860
gcttccgc ttccctcgctc actgactcg tgcgtcggt cggtcgctg cggcgacgg	7920
tatcagctca ctcaaaaggcg gtaatacggt tatccacaga atcagggat aacgcagaa	7980
agaacatgtg agcaaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgcgt	8040
cgttttcca taggctccgc cccccctgacg agcatcacaat aaatcgacgc tcaagtca	8100
ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg	8160
tgcgtctcc tggccgacc ctggccctta ccggatactt gtccgcctt ctccctcgg	8220
gaagcgtggc gctttctcat agtcacgt gtaggtatct cagttcggtg taggtcggt	8280
gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctc gccttattccg	8340
gtaactatcg tctttagtcc aacccggtaa gacacgactt atcggcaactg gcagcagcc	8400
ctggtaacag gattacgaga gcgaggatgt taggcgtgc tacagagttc ttgaagtgg	8460
ggcctaacta cgctacact agaagaacag tatttggat ctgcgtctc ctgaagccag	8520
ttaccttcgg aaaaagagtt ggtagcttt gatccggcaa acaaaccacc gctggtagcg	8580
gtgggtttt tgggtcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc	8640
ctttgatctt ttctacgggg tctgacgctc agtggacaca aaactcacgt taaggattt	8700
tggcatgag attatcaaaa aggatctca cctagatctt tttaaattaa aaatgaagtt	8760
ttaaatcaat ctaaaagtata tatgataaa ctgggtctga cagttacca tgcttaatca	8820
gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagtgcc tgactcccc	8880
tcgtgttagat aactacgata cgggagggtt taccatctgg ccccaagtgc gcaatgata	8940
cgcgagaccc acgctcaccc gctccagatt tatcagcaat aaaccagcca gcccggaaagg	9000
ccgagcgcag aagtggctt gcaactttt ccgcctccat ccagcttatt aatttgttgc	9060
gggaagactag agtaagtagt tcgcccgtt atagttgcg caacgttgc gccattgtca	9120
caggcatcggt ggtgtcacgc tcgtcggtt gatggcttc attcaagtcg ggttcccaac	9180
gatcaaggcg attacatga tccccatgt tgcggaaaa agcggttagc tccttcggc	9240
ctccgtatcg tgcgtatgtt aagttggccg cagtgttattc actcatgtt atggcagcac	9300
tgcataattc tcttactgtc atgcccattt taagatgtt ttctgtgact ggtgagtagt	9360
caaccaagtc attctgagaa tagtgtatgc ggccgaccgg ttgccttgc ccggcgtcaa	9420
tacgggataa taccgcgcca catagcagaa cttaaaagt gtcatcatt gggaaacgtt	9480
cttcggggcg aaaactctca aggatcttac cgctgttgcg atccagttcg atgtacccca	9540
ctcggtcacc caactgtatc tcagcatctt ttacttccat cagcgtttt gggtgagcaa	9600
aaacagggaaag gcaaaatgcc gcaaaaaagg gaataaggc gacacggaaa tggtgaatac	9660
tcatactttt ctttttcaatttatttgcgatccat ggttattgt ctcgtatcg	9720
gatacatatt tgaatgtatt tagaaaaata aacaaatagg gttccgcgc acatccccc	9780
ggaaaggcgcc acgtgtacgtc taagaaacca ttattatcat gacattaacc tataaaaata	9840
ggcgtatcac gaggccctt cgtctcgcc gttcggtga tgacggtaaa aacctctgac	9900
acatgcagct cccggagacg gtcacagctt gtctgtatcg ggtgcgggg agcagacaag	9960
cccggtcaggc cgcgtcagcg ggtgttggcg ggtgtcgccc ctggcttaac tatgcggcat	10020

## GC830-PCT.txt

cagagcagat	tgtactgaga	gtgcaccata	aaattgtaaa	cgttaatatt	ttgttaaaat	10080
tcgcgttaaa	ttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcgcca	10140
tcccttataa	atcaaaaagaa	tagcccgaga	taggggttag	tgttgttcca	gtttggaa	10200
agagtccact	attaaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	10260
gcgatggccc	actacgtgaa	ccatcaccca	aatcaagtt	tttggggtcg	aggtgccgta	10320
aagcaactaa	tcgaaaccct	aaagggagcc	cccgatttag	agcttgacgg	ggaaagccgg	10380
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaaggagc	gggcgtctagg	gchgctggca	10440
gtgttagcggt	cacgctgcgc	gtaaccacca	caccgcgc	gcttaatgcg	ccgctacagg	10500
gcgcgtacta	tgttgtctt	gacgtatgcg	gtgtgaaata	ccgcacagat	gchtaaggag	10560
aaaataccgc	atcaggcgcc	attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	10620
ggtgcgggccc	tcttcgctat	tacgcccagct	ggcggaaagg	ggatgtgctg	caaggcgatt	10680
aagttgggta	acgcccagggt	tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgccc	10739

<210> 5  
<211> 2169  
<212> DNA  
<213> Aspergillus awamori var. kawachi

<400> 5						
atgtcgttcc	gatctttct	cgcctgagc	ggccttgtct	gctcggtt	ggcaagtgt	60
atttccaagc	gcgcgacctt	ggattcgtgg	ttgagcaacg	aagcaccgt	ggcccgta	120
gcgatcctga	ataaacatcg	ggcggacgg	gcttgggtgt	cgggcgcg	ctctggcatt	180
gtcggttcca	gtccccagcac	cgataacccg	gactgtatgt	ttttagttcg	gattatgaat	240
gtgtctgtgt	tgattgtatgc	tgactggcg	gtctttgtat	gattttagac	ttctacacct	300
ggactcgcga	ctctggtctc	gtcatcaaga	ccctcgctga	ccttccgc	aatggagata	360
ctgatctcct	ttccaccatt	gagcactaca	tctctctca	ggcaattatt	cagggtgtca	420
gtaacccttc	tggtagtctg	tccagcggt	gtcttgggt	gcccaagttc	aatgtcgatg	480
agactgccta	caccggttct	tggggacggc	cgcagcgtg	tggctctg	ctgagagcaa	540
ctgctatgtat	cggcttggg	cagtggctgc	ttgtatgttc	tccacctcct	tgctgtctgat	600
ctgcaacata	tgttagccgac	tggtagggac	aatggctaca	ccagcgctgc	aacagagatt	660
gtttggccccc	tctttaggaa	cgacctgtcg	tatgtggctc	agtaactggaa	ccagacggga	720
tatgtgtgt	ttgattgtatc	gggggtcaag	ggtgtttgt	catcgagct	aacttcgcgg	780
tcgcagatct	ctgggaagaa	gttaatggct	cttccttctt	cactattg	gtgcaacacc	840
gcgcctcggt	cgaaggtagt	gccttcgcga	cgcccgctgg	ctcgctctgc	tcctgggtgt	900
attcgcaaggc	acccatcgatt	ctctgttact	tgcagtcctt	ctggaccggc	agctacatcc	960
tggccaactt	tgacagcagc	cgttccggca	aggacacaaa	caccctcctg	ggaagcatcc	1020
acaccttta	tcctgaggct	ggatgcgacg	actccacett	ccagccctgc	tcccgcg	1080
cgcctccaa	ccataaggag	gtttagact	cttcctcg	gatctatact	ctcaacgat	1140
gtctcagtga	cagtggcg	gttgcgg	gtcggtagcc	tgaggatagc	tactacaacg	1200
gcaaccctg	gttcctgtgc	accttggctg	ccgcggaa	gctgtacgt	gctctgtacc	1260
agtgggacaa	gcaggggctg	ttggagatca	cagacgtgc	acttgacttc	ttcaaggctc	1320
tgtacagtgg	tgtgtccacc	ggcacgtact	cttcgtccag	ctcgacctat	agcagcattg	1380
ttagtggcg	caagactttc	gctgtatgg	ttgtttctat	tgtgttaat	ctacgctaga	1440
cgagcgtca	tatccacaga	gggtgcgtac	taacaggatt	aggaaactca	cgccgca	1500
aacggctctc	tgtctgagca	attcgacaa	tctgtatggcg	acgagcttc	tgctcgcgat	1560
ctgacctgg	cttacgctgc	tctgtcgacc	gccaacaacc	gtcgtaattc	tgtcg	1620
ccgtcttggg	gtgagacacc	tgccagc	gtgcccggca	cctgtcg	tacccctgt	1680
tctgttaccc	acagcgtgt	gaccgtcacc	tcgtggccg	gcatcg	tactgg	1740
accactacga	cgcgtactac	cactggatcg	ggcgccgtg	cctcgacc	caagaccacc	1800
acaactgcta	gtaagacc	caccactacg	tcctcgac	cctgcacc	ccccactg	1860
gtagctgtga	cctttgatct	gacggcgacc	accacctacg	gchagaacat	ctacctgg	1920
gggtcgatct	ctcagctcg	tgactggag	accagcgat	gcatagct	gagcg	1980
aagtacactt	ccagcaaccc	gctttgtat	gtaactgt	ctctggc	tggtaggt	2040
ttttagtaca	attcatccg	cgtcgagagc	gtgactccg	tggagtg	gagcgaccc	2100
aaccggaaat	acaccgttcc	tcaggcgtgc	ggcgagtcg	ccgcgacgg	gaccgacacc	2160
tggcggtag						2169

<210> 6  
<211> 24

## GC830-PCT.txt

<212> PRT  
<213> Aspergillus awamori var. kawachi

<400> 6  
Met Ser Phe Arg Ser Leu Leu Ala Leu Ser Gly Leu Val Cys Ser Gly  
1 5 10 15  
Leu Ala Ser Val Ile Ser Lys Arg  
20

<210> 7  
<211> 615  
<212> PRT  
<213> Aspergillus awamori var. kawachi

<400> 7  
Ala Thr Leu Asp Ser Trp Leu Ser Asn Glu Ala Thr Val Ala Arg Thr  
1 5 10 15  
Ala Ile Leu Asn Asn Ile Gly Ala Asp Gly Ala Trp Val Ser Gly Ala  
20 25 30  
Asp Ser Gly Ile Val Val Ala Ser Pro Ser Thr Asp Asn Pro Asp Tyr  
35 40 45  
Phe Tyr Thr Trp Thr Arg Asp Ser Gly Leu Val Ile Lys Thr Leu Val  
50 55 60  
Asp Leu Phe Arg Asn Gly Asp Thr Asp Leu Leu Ser Thr Ile Glu His  
65 70 75 80  
Tyr Ile Ser Ser Gln Ala Ile Ile Gln Gly Val Ser Asn Pro Ser Gly  
85 90 95  
Asp Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu  
100 105 110  
Thr Ala Tyr Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala  
115 120 125  
Leu Arg Ala Thr Ala Met Ile Gly Phe Gly Gln Trp Leu Leu Asp Asn  
130 135 140  
Gly Tyr Thr Ser Ala Ala Thr Glu Ile Val Trp Pro Leu Val Arg Asn  
145 150 155 160  
Asp Leu Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp Leu  
165 170 175  
Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Ile Ala Val Gln His  
180 185 190  
Arg Ala Leu Val Glu Gly Ser Ala Phe Ala Thr Ala Val Gly Ser Ser  
195 200 205  
Cys Ser Trp Cys Asp Ser Gln Ala Pro Gln Ile Leu Cys Tyr Leu Gln  
210 215 220  
Ser Phe Trp Thr Gly Ser Tyr Ile Leu Ala Asn Phe Asp Ser Ser Arg  
225 230 235 240  
Ser Gly Lys Asp Thr Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp  
245 250 255  
Pro Glu Ala Gly Cys Asp Asp Ser Thr Phe Gln Pro Cys Ser Pro Arg  
260 265 270  
Ala Leu Ala Asn His Lys Glu Val Val Asp Ser Phe Arg Ser Ile Tyr  
275 280 285  
Thr Leu Asn Asp Gly Leu Ser Asp Ser Glu Ala Val Ala Val Gly Arg  
290 295 300  
Tyr Pro Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Cys Thr  
305 310 315 320  
Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Gln Trp Asp Lys  
325 330 335  
Gln Gly Ser Leu Glu Ile Thr Asp Val Ser Leu Asp Phe Phe Lys Ala  
340 345 350

GC830-PCT.txt

Leu Tyr Ser Gly Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Thr  
355 360 365  
Tyr Ser Ser Ile Val Ser Ala Val Lys Thr Phe Ala Asp Gly Phe Val  
370 375 380  
Ser Ile Val Glu Thr His Ala Ala Ser Asn Gly Ser Leu Ser Glu Gln  
385 390 395 400  
Phe Asp Lys Ser Asp Gly Asp Glu Leu Ser Ala Arg Asp Leu Thr Trp  
405 410 415  
Ser Tyr Ala Ala Leu Leu Thr Ala Asn Asn Arg Arg Asn Ser Val Val  
420 425 430  
Pro Pro Ser Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Thr Cys  
435 440 445  
Ala Ala Thr Ser Ala Ser Gly Thr Tyr Ser Ser Val Thr Val Thr Ser  
450 455 460  
Trp Pro Ser Ile Val Ala Thr Gly Gly Thr Thr Thr Ala Thr Thr  
465 470 475 480  
Thr Gly Ser Gly Gly Val Thr Ser Thr Ser Lys Thr Thr Thr Ala  
485 490 495  
Ser Lys Thr Ser Thr Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr  
500 505 510  
Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr Thr Tyr Gly Glu  
515 520 525  
Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr  
530 535 540  
Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asn Pro  
545 550 555 560  
Leu Trp Tyr Val Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr  
565 570 575  
Lys Phe Ile Arg Val Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp  
580 585 590  
Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys Gly Glu Ser Thr Ala  
595 600 605  
Thr Val Thr Asp Thr Trp Arg  
610 615

<210> 8  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer

<400> 8  
caacatgcat accttctcca agtcctc

28

<210> 9  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer

<400> 9  
ttaacgccac gaatcattca ccgtc

25

<210> 10  
<211> 25

GC830-PCT.txt

<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer

<400> 10  
caccatgtcg ttccgatctc ttctc

25

<210> 11  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer

<400> 11  
ctaccgccag gtgtcggtca c

21